MrPaw

Ryan Giordano (rgiordano@berkeley.edu, UC Berkeley), ..,

Berkeley BSTARS September 2025

Are US non-voters becoming more Republican?

Blue Rose research says yes:

"Politically disengaged voters have become much more Republican, And because less-engaged voters swung away from [Democrats], an expanded electorate meant a more Republican electorate."

[Blue Rose Research, 2024] (On Ezra Klein show, major professional pollsters)

***On Data and Democracy* says no:**

"Claims of a decisive pro-Republican shift among the overall non-voting population are not supported by the most reliable, large-scale post-election data currently available."

[Bonica et al., 2025] (Berkeley professor co–author, major professional researchers)

Several factors drive the disagreement:

- The problem is very hard (it's difficult to poll non–voters)
- Different data sources
  - Blue Rose aggregates its own private data
  - The *On Data and Democracy* posts use public data, e.g. the cooperative election study (CES).
- **Very different statistical methods:** ⋆
  - Blue Rose uses Bayesian hierarchical modeling
  - The CES uses calibration weighting

Our work won't resolve the dispute. (Anyway, we'd need access to Blue Rose's private data and modeling to even try.)

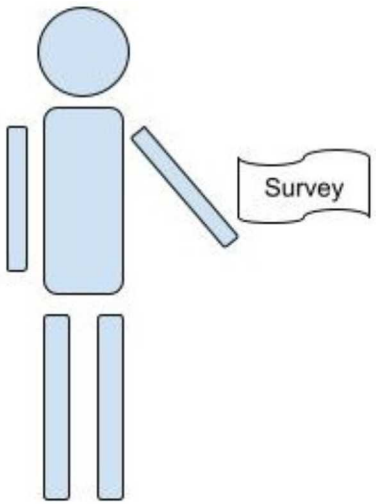
But we can form a like–to–like comparison between the methodologies. (And hope that Blue Rose tries our software package.)

The basic problem

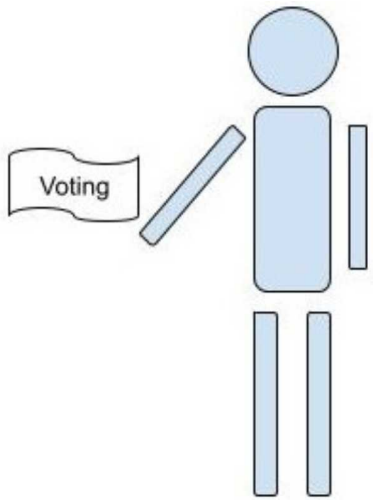
We have a survey population, for whom we observe:

- Covariates x (e.g. race, gender, zip code, age, education level)
- Responses y (e.g. A binary response to “do you support policy such–and–such”)

We want the average response in a target population, in which we observe only covariates.



Observe (x_s, y_s) for $s = 1, \ldots, \mathcal{N}_S$



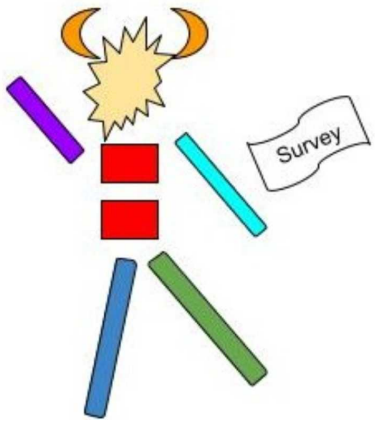
Observe x_p for $p = 1, \ldots, \mathcal{N}_T$

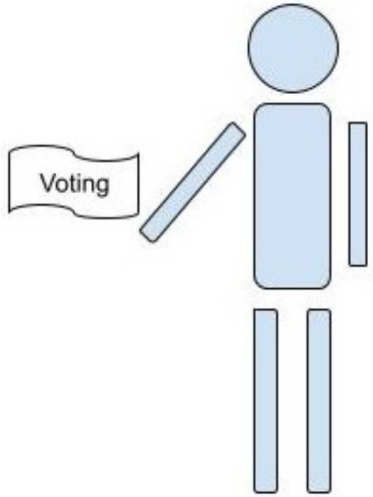
The basic problem

We have a survey population, for whom we observe:

- Covariates x (e.g. race, gender, zip code, age, education level)
- Responses y (e.g. A binary response to “do you support policy such–and–such”)

We want the average response in a target population, in which we observe only covariates.





Observe (x_s, y_s) for $s = 1, \ldots, \mathcal{N}_S$

Observe x_p for $p = 1, \ldots, \mathcal{N}_T$

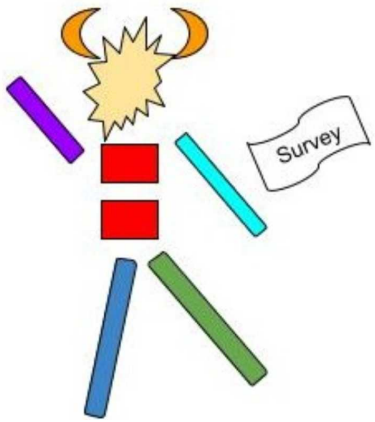
The problem is that the populations are very different.

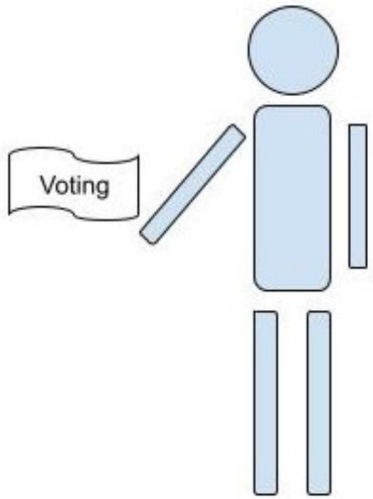
The basic problem

We have a survey population, for whom we observe:

- Covariates x (e.g. race, gender, zip code, age, education level)
- Responses y (e.g. A binary response to “do you support policy such–and–such”)

We want the average response in a target population, in which we observe only covariates.





Observe (x_s, y_s) for $s = 1, \ldots, \mathcal{N}_S$

Observe x_p for $p = 1, \ldots, \mathcal{N}_T$

The problem is that the populations are very different.

Our survey results may be biased.

How can we use the covariates to say something about the target responses?

Two approaches

We want $\mu := \frac{1}{N_T} \sum_{n \in \mathcal{N}_T} y_p$, but don't observe y_p in the target population.

- Assume $p(y|x)$ is the same in both populations,
- But the distribution of x may be different in the survey and target.

Calibration weighting

$\hat{\mu}_{\text{CAL}} := \frac{1}{N_S} \sum_{n \in \mathcal{N}_S} w_s y_s$

For some weights w_s on the survey observations.

Model–based

References

Blue Rose Research. 2024 Election Retrospective Presentation. https://data.blueroseresearch.org/2024retro-download, 2024. Accessed on 2024-10-26.

A. Bonica, R. Fordham, J. Grumbach, and E. Tiburcio. Did non-voters really flip Republican in 2024? The evidence says no. https://data4democracy.substack.com/p/did-non-voters-really-flip-republican, April 2025.